

Traitement statistique du signal et des images

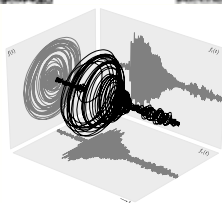
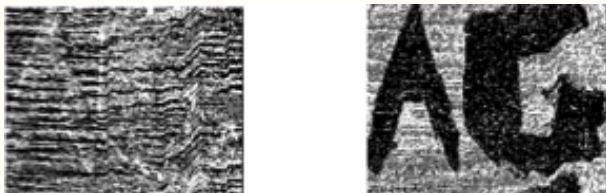


2017

- ▶ Signaux, Modèles et Applications, **19 personnes**
- ▶ 13 permanents : 3 EC-Lille, 2 CNRS, 4 Univ. Lille, 4 Telecom-lille
- ▶ 6 doctorants
- ▶ Objet : traitement du signal et des images, apprentissage statistique (aide à l'analyse, à la décision)

- ▶ **Signaux** : provenant de capteurs (bases de photos, sons, imagerie, vidéos, capteurs de pollution, de turbulences, ...)
- ▶ **Modèles** :
 - physiques et statistiques : savoir modéliser des signaux
 - paramétriques : loi de probabilité
 - non paramétriques : histogramme, dictionnaires
 - évidentiels : fusion de capteurs multi-modaux
 - informationnels : sécurité et fuite d'information
- ▶ **Applications** (exemples)
 - **sécurité** : détection de signaux cachés, d'intrusion
 - **vidéo** : suivis d'objects, vidéo-surveillance
 - **environnement** : localisation de sources de pollutions
 - **physique** : fusion de mesures (écoulement fluide)
 - **imagerie** : biologique, polarimétrique, ...
 - ...

- ▶ **Imagerie polarimétrique** [V. Devlaminck,
 - geometry of information, polarized signals, experimental devices

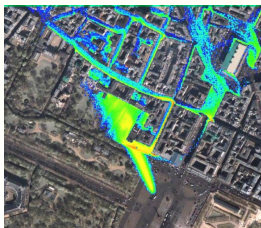


→ thèse de Julien Flamant

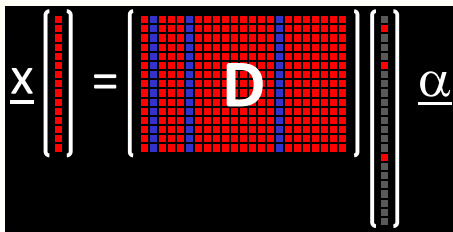
► Méthodes probabilistes & approches bayésiennes

- Bayesian estimation, probabilistic models, dictionary learning, data fusion, belief functions

Localisation de sources de pollutions (collab. ARIA Tech.)



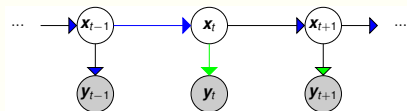
Dictionnaires et pb inverses

$$\underline{x} = \underline{D} \underline{\alpha}$$
The equation $\underline{x} = \underline{D} \underline{\alpha}$ is displayed on a black background. The vector \underline{x} is represented by a vertical column of red squares. The matrix \underline{D} is represented by a grid of red and blue squares, with a white letter 'D' in the center. The vector $\underline{\alpha}$ is represented by a vertical column of grey squares, with a few red squares at the bottom.

Inference and Estimation in State-Space Models

► Let us consider :

- a set of hidden states $\mathbf{x}_t \in \mathbb{R}^{d_x}$, $t = 1, \dots, T$
- a set of observations $\mathbf{y}_t \in \mathbb{R}^{d_y}$, $t = 1, \dots, T$



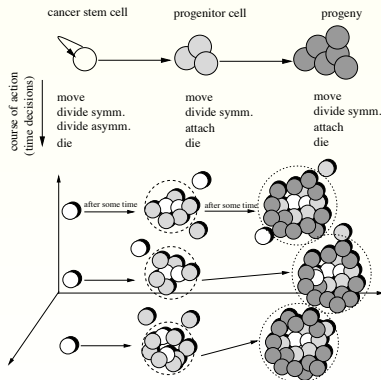
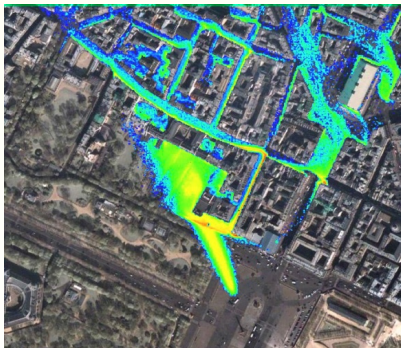
- $p_{\theta}(\mathbf{x}_t | \mathbf{x}_{t-1}) \Rightarrow$ the **hidden** state evolves dynamically (Markov)
- $p_{\theta}(\mathbf{y}_t | \mathbf{x}_t) \Rightarrow$ **observations** are related to the state
- Both distributions are parametrized by θ

Goal : Inference of the hidden state $p(\mathbf{x}_t | y_{1:t})$ & estimation of θ

- In real-life problems, the models are complex and $p(\mathbf{x}_t | y_{1:t})$ cannot be computed analytically
- \rightsquigarrow **Monte Carlo** methods : $p(\mathbf{x}_t | y_{1:t})$ is approximated by sampling **random numbers**

Inference and Estimation in State-Space Models

- ▶ Problems in (so) many disciplines :
 - Tracking in radar, space debris, computer vision, etc
 - Prediction of economical data, biological populations, etc
 - Weather forecasting



Stage possible en 2017 5-6 mois

CV+Lettre de motivation ↦ F. Septier, C. Elvira

Inference and Estimation in State-Space Models

- ▶ Problems in (so) many disciplines :
 - Tracking in radar, space debris, computer vision, etc
 - Prediction of economical data, biological populations, etc
 - Weather forecasting

↔ **Main challenge :**

- Design good samplers of random numbers in complex/high-dimensional problems

⇒ **Research objectives :**

- **Adaptive** sampling schemes - *learn from the past...*
- Efficient dynamic allocation of the computational effort

↔ Application : real data from biological system or space debris...

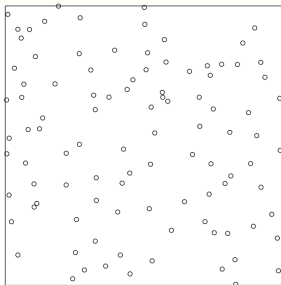
Stage possible en 2017 5-6 mois

CV+Lettre de motivation ↦ F. Septier, C. Elvira

► Monte Carlo methods

- algorithms for inference,
- DPP = distributions de points “qui se repoussent”
- ex : particules chargées, arbres d’une forêt, cellules d’un tissu, avions au-dessus d’un aéroport...

⇒ méthodes BNP basées sur des **DPP**...

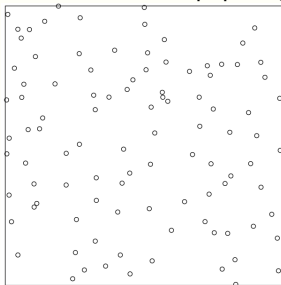


Sujet M2/thèse : *Adaptive subspace discovery using determinantal point processes for signal processing*

Encadrants :

Pierre Chainais, MCF HDR École Centrale, Équipe SigMA

Rémi Bardenet, chercheur CR2 CNRS, Équipe SigMA



Problèmes inverses en traitement d'image

Image d'origine en niveaux de gris



Problèmes inverses en traitement d'image

Image bruitée \implies débruitage ?



Problèmes inverses en traitement d'image

Image bougée \Rightarrow déconvolution ?



Inpainting combinant ondelettes et cosinus locaux (MCA)



Inpainting combinant ondelettes et cosinus locaux (MCA)



Problèmes inverses : mal posés

Idée : chercher la solution dans un espace restreint



utiliser des a priori bien choisis



utiliser une représentation adaptée

► **bases de fonctions = dictionnaires**

- exponentielles complexes : Fourier
⇒ *pour les signaux périodiques ou stationnaires*
- cosinus par blocs ⇒ *format JPEG*
- cosinus locaux : signaux localement stationnaires
⇒ *dans votre lecteur MP3!*

► **apprentissage de dictionnaires** : à partir des données

→ thèse de Hong Phuong Dang

Dictionnaire appris sur une image

Apprentissage statistique

- ▶ découper l'image en blocs appelés *patches*, par ex. 7×7 ,
- ▶ identifier des blocs typiquement représentatifs
- ▶ s'assurer que la représentation utilisera peu de motifs (*parcimonie ou sparsity*)



Inverse problems in image processing

Denoising



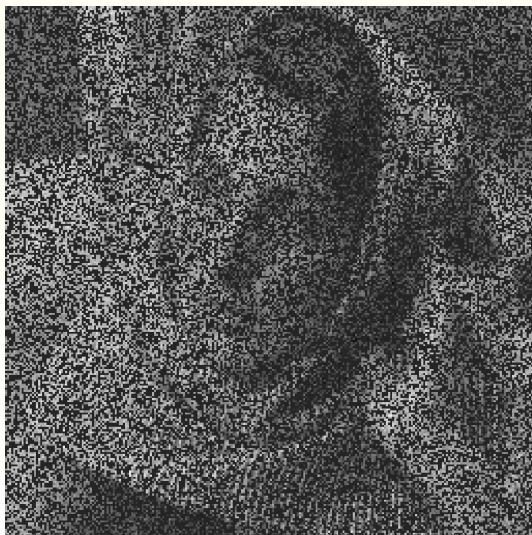
Inverse problems in image processing

Denoising



Inverse problems in image processing

Inpainting



Inverse problems in image processing

Inpainting



Stéganographie & Stéganalyse

► Stéganographie et stéganalyse



Stéganographie



Stéganalyse

- theory of communications, machine learning, watermarking, steganography/steganalysis, authentication

Stéganographie & Stéganalyse

► Stéganographie :

l'art de **cacher** un message dans un support

- sélectionner des “cachettes”
(contours, textures...)
- modifier le support de façon “imperceptible”
(modifier le dernier bit seulement...)

► Stéganalyse :

l'art de **détecter** la présence d'un message caché

- calculer des “caractéristiques sensibles”,
- apprendre à distinguer images *stego* / *cover*

Stéganographie & Stéganalyse

Exemple 1

Image d'origine



Stéganographie & Stéganalyse

Exemple 1

Cachettes possibles



Stéganographie & Stéganalyse

Exemple 1

Pixels modifiés



Stéganographie & Stéganalyse

Exemple 1

Image modifiée



Stéganographie & Stéganalyse

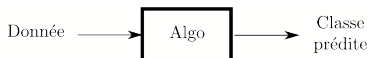
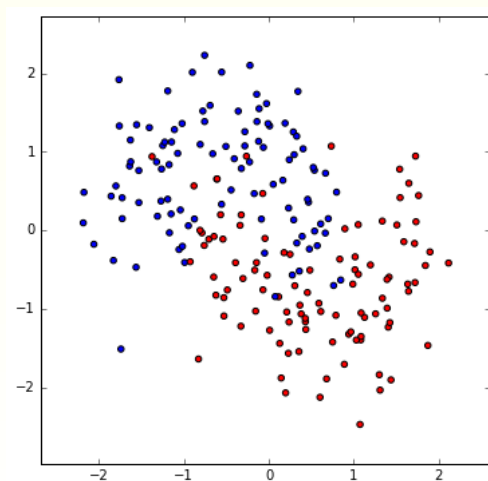
Exemple 1

Image d'origine



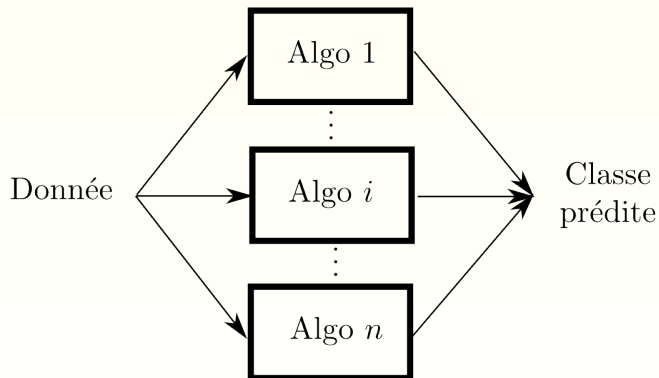
Fusion de données

Prenons un problème de **classification** pour les données suivantes :



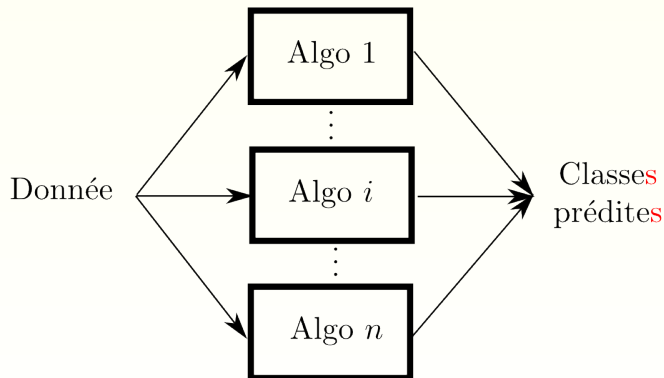
Fusion de données

Admettons que je sois **riche** et j'ai un serveur de calcul. Du coup j'entraîne **plusieurs** algorithmes !



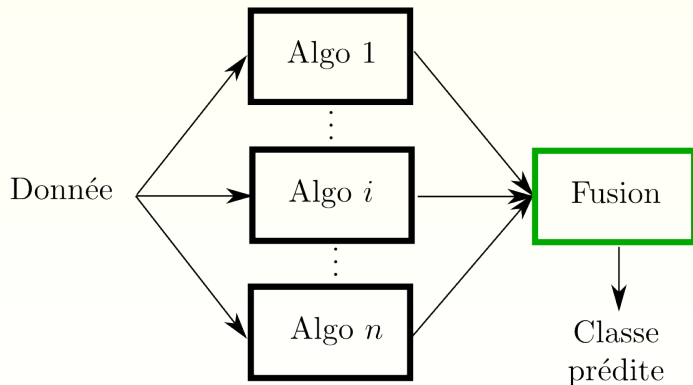
Fusion de données

Admettons que je sois **riche** et j'ai un serveur de calcul. Du coup j'entraîne **plusieurs** algorithmes !



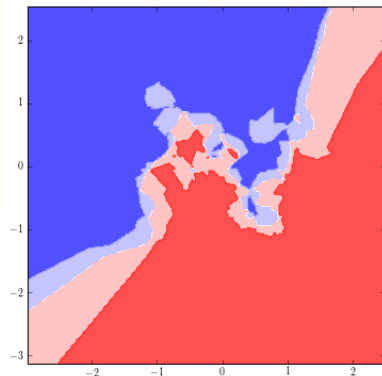
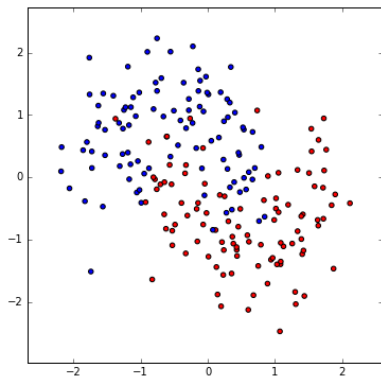
Fusion de données

Admettons que je sois **riche** et j'ai un serveur de calcul. Du coup j'entraîne **plusieurs** algorithmes !



Fusion de données

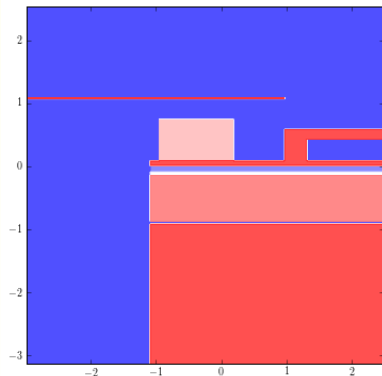
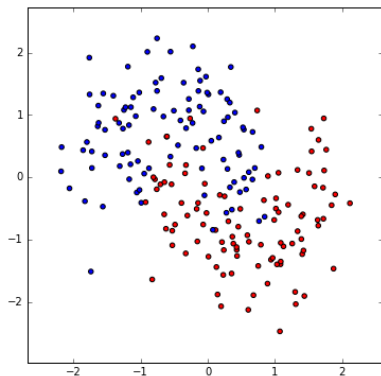
Sur mon dataset : un 1^{er} algo donne :



(k-PPV)

Fusion de données

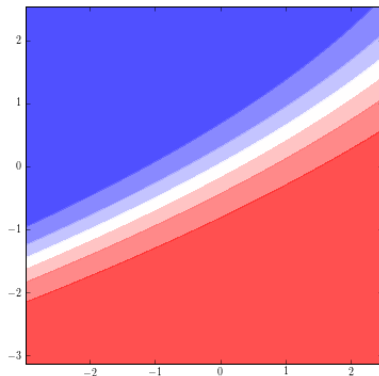
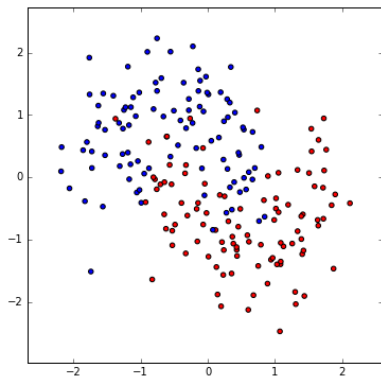
Sur mon dataset : un 2^{ème} algo donne :



(Arbre de décision)

Fusion de données

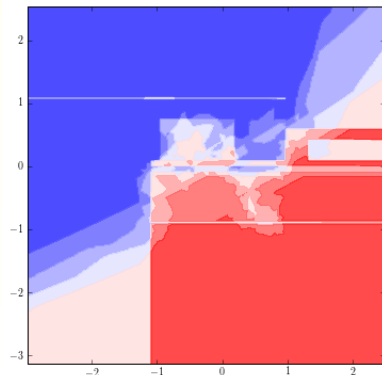
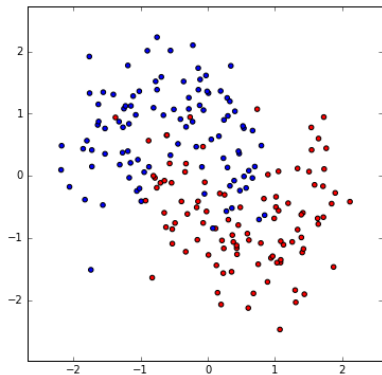
Sur mon dataset : un 3^{ème} algo donne :



(naïf Bayésien)

Fusion de données

Sur mon dataset : leur fusion donne :



(Vote majoritaire)

Fusion de données

Est-ce vraiment **utile** ?

- 1 Sur ce jeu de données oui ! Le meilleur des 3 algorithmes a 85% de bonnes réponses, le **vote majoritaire** monte à 88%.
- 2 Peut-on faire **mieux** qu'un vote majoritaire ?
→ vers des algos de fusion **robustes** et **adaptatifs**.
- 3 Ca peut vous rendre **riche**

The Netflix logo, consisting of the word "NETFLIX" in white, bold, sans-serif capital letters with a black drop shadow, centered on a solid red rectangular background.

NETFLIX

\$1,000,000